

**GUIDANCE FOR BENCHMARK DOSE (BMD) APPROACH -
CONTINUOUS DATA**

DPR MT-2

**Health Assessment Section
Medical Toxicology Branch**

**Department of Pesticide Regulation
September 2004**

ACKNOWLEDGEMENT

This Document is a product of the Health Assessment Section (HAS) in Medical Toxicology Branch, Department of Pesticide Regulation, Cal/EPA.

This document should be cited as:

*DPR MT-2, 2004. Guidance for Benchmark Dose (BMD) Approach - Continuous data.
Medical Toxicology Branch, Department of Pesticide Regulation, California
Environmental Protection Agency, Sacramento, CA.*

The following staff (alphabetically listed) contributed to the production of this document:
C. Jenks, Ph.D., E. Kwok, Ph.D., L. Lim, Ph.D., D.A.B.T., and N. R. Reed, Ph.D., D.A.B.T.
(Lead person).

LIST OF ABBREVIATIONS

σ	standard deviation
AIC	Akaike's Information Criterion
ALT	alanine aminotransferase
BMD	Benchmark Dose
BMR	Benchmark Dose Response
Cal/EPA	California Environmental Protection Agency
ChE	cholinesterase
DPR	Department of Pesticide Regulation
ED	Effective Dose at a specified response level (e.g., ED ₀₅ : Effective Dose at 5% response or BMR); also referred to as BMD (e.g., BMD ₀₅)
HAS	Health Assessment Section of Medical Toxicology Branch, DPR
LED	Lower bound of ED (e.g., LED ₀₅ : lower 95 th confidence bound of ED ₀₅); also referred to as BMDL (e.g., BMDL ₀₅)
LOEL	Lowest-Observed-Effect Level
MCV	mean corpuscular volume
MOE	margin of exposure
MTD	maximum tolerated dose
NOEL	No-Observed-Effect Level
OP	organophosphate
RBC	red blood cell
PB/PK model	physiologically based pharmacokinetic model
RPF	relative potency factor
USEPA	United State Environmental Protection Agency

I. INTRODUCTION

This document provides the necessary background and guidance for a consistent application of the benchmark dose (BMD) approach in the dose-response assessment of continuous data. It does not include guidance for the analysis of “nested data” that are most commonly seen in reproductive and developmental endpoints when the response of fetuses from one litter are inter-related. Crucial scientific issues in this document underwent a series of discussions and deliberation within the Health Assessment Section (HAS) of Medical Toxicology Branch to ensure sound scientific considerations. The guidance for quantal data analysis is available in a parallel document, *Guidance for Benchmark Dose (BMD) Approach - Quantal Data* (DPR MT-1, 2004).

The current risk assessment practice assumes that a threshold dose exists for effects other than oncogenicity, i.e., toxicologically significant effects are not likely to occur below the threshold dose. Two approaches can be used to define this threshold dose. The traditional approach determines the toxicity threshold as the no-observed-effect level (NOEL). NOEL is the highest dose in a study at which no effects are established (i.e., observed or measured). The next higher dose at which effects are seen is the lowest-observed-effect Level (LOEL). These NOELs and LOELs may be established based on statistically significant responses (e.g., $p \leq 0.05$) at the LOEL or by the evidence of a continuum of response with increasing dose. In this approach, the determination of the threshold dose is dictated by the dose selection in a toxicity study.

An alternative to the NOEL-LOEL approach is the BMD approach. It involves fitting a mathematical model to the entire dose-response dataset for an endpoint, and allowing the model to estimate the threshold dose corresponding to a level of benchmark response (BMR). This BMR is set at a certain level (e.g., 1%, 5%, 10%) as defined by the risk assessor. The BMD is either the model's best estimate of the effective dose (ED) at the BMR or the statistical 95th percent lower bound of ED (LED). Accordingly, BMD can be expressed as ED₀₁, ED₀₅, ED₁₀, or LED₀₁, LED₀₅, LED₁₀. Other comparable terms have also been used, such as BMD₀₁, BMD₀₅, BMD₁₀, and BMDL₀₁, BMDL₀₅, BMDL₁₀.

The NOEL-LOEL approach is relatively simple in that they can be determined directly from a study. Conversely, the BMD approach requires an extra step of fitting models to the dose-response data before determining the EDs and LEDs. However, the NOEL-LOEL approach has several limitations. It tends to focus only on data points at the apparent NOEL and LOEL, and not making full use of the entire dataset. This could result in different NOELs for an endpoint from two "identical" studies that differ only by the choice of dose for study. The NOEL-LOEL approach also tends to "reward" studies with smaller sample size or greater variations in endpoint measurement by assigning a higher NOEL based on statistical comparison to the controls. In reality, data from this type of study could mean greater uncertainty and higher probability for a false negative. Moreover, when a NOEL cannot be directly determined from a study (e.g., effects are present at the lowest tested dose), the NOEL-LOEL approach is inadequate to define a threshold for risk assessment. The current default practice is to divide the LOEL by a somewhat arbitrary uncertainty factor, usually within 10 (e.g., 1, 3, 10). Given the same dataset, the BMD approach can overcome these limitations of the NOEL-LOEL approach.

The advantages of the BMD approach are summarized below:

- Characterize the dose-response curve by using all pertinent data points
- Allow consistency in establishing the threshold dose from all studies and chemicals (i.e., corresponding to a given response level for an endpoint)
- Account for the greater uncertainty due to smaller sample size or greater variation in endpoint measurements or observations when the threshold dose is established as the LED
- Consistently estimate the threshold dose when no NOEL can be established (i.e., the lowest tested dose is the LOEL)

Although the BMD approach was introduced in the 90's, it has not been widely used until recently. One of the reasons is that suitable mathematical models for the variety of data types were limited and costly. This obstacle was removed recently with the BMD software made publicly available by USEPA in 2001 and subsequently updated. The other reason was the need for consistent criteria in the application of BMD approach. These include: the choice of model, the criteria for use of data in modeling, and the choice of BMR and BMD for risk assessment. The guidance for these areas is provided in this document.

II. GENERAL GUIDANCE

Prior to applying the benchmark dose (BMD) modeling, all toxicity endpoints should be identified in the Hazard Identification phase and accompanied with the identification of the no-observed-effect level (NOELs) (if possible) and the lowest-observed-effect level (LOEL).

To make the full use of a BMD approach and consistently account for the differential qualities between studies (i.e., sample size, measurement or observational variations), the HAS consensus is to identify the risk assessment threshold dose as the lower bound of the effective dose (LED). Nevertheless, a presentation of both the ED (best estimate of the effective dose) and LED values in the dose-response assessment is advisable as they provide more thorough information on modeling. For uniformity, the DPR convention for the threshold dose is expressed in "LED" (e.g., LED₀₁, LED₀₅, LED₁₀) and not BMDL (lower bound of BMD, e.g., BMDL₀₁, BMDL₀₅, BMDL₁₀).

III. SUMMARY GUIDANCE

The HAS guidance is summarized below. Detail discussions for the modeling process pertinent to continuous data are presented in Section IV.

- 1. Endpoint Selection :** Considerations are given to the effects which are toxicological significant and/or adverse. Sensitive and biologically relevant endpoints (e.g. activities of serum alanine aminotransferase, ALT) can also be identified for BMD modeling. Data for several pertinent endpoints should be modeled to ensure finding the lowest BMD from all datasets. Alternatively, BMD approach can be applied on a per need basis (e.g., when no NOEL can be established).

2. **Data Criteria :** The modeling requires data of individual test subjects or their summarized form (group dose, size, mean, standard deviation σ). A dataset should have at least two treated groups¹ other than the control, with either a significant change in response with increasing dose² (positive trend; $p \leq 0.05$) or a significant pair-wise increase in response ($p \leq 0.05$) in at least one treated group. Datasets with near maximum response at the lowest tested dose is generally not a good candidate because of the extensive extrapolation (see below).
3. **Uncertainties in Extrapolation :** Extensive extrapolation below the experimental range of response should be avoided because it tends to introduce greater uncertainties.
4. **Data Conversion :** Theoretically, individual data are preferred for the modeling as they allow the model to fully utilize the variable distribution for defining the LED. The benchmark response (BMR) is commonly expressed in terms of changes in group means relative to the control (e.g., 5% reduction of control mean body weight). However, a given level of BMR does not necessarily carry the same biological significance for all endpoints (e.g., 10% change in mean body weight versus 10% change in liver enzyme). A possible "comparable metric" is to express the BMR as changes in a proportion of the population affected. By dichotomizing the continuous data (i.e., dividing the population into the "affected" and "unaffected"), a comparison of endpoint sensitivity based on the BMD may be possible through defining the BMR at a fixed proportion of population affected. Detailed discussions on this approach are presented in Section IV, Step 4 and Step 6.
5. **Choice of Model and Options:** When more than one model can adequately describe a dataset, the model with the best fit should generally be used. The model fit criteria and considerations for model selection are given in Section IV, Step 5. If none of the available models can fit, or the model fit is poor in the region near the BMR (e.g., disparity between the "observed" and the "estimated"), data point(s) high above the BMR may be excluded, while still retaining the minimum number of dose group needed for modeling (see Section IV, Step 2). After eliminating a dose group, the variance homogeneity pattern may be altered and should be re-evaluated (e.g., change from non-homogeneous to homogeneous). Physiologically based pharmacokinetic (PB/PK) model can also improve the model fit by estimating the dose or concentrations of the parent chemical or its active metabolites at the target site(s). When no model can adequately describe the dose response relationship, it may be necessary to revert to the NOEL-LOEL approach.
6. **Define BMR :** For characterizing the risk, the HAS default is to define the BMR based on a percentage of "relative deviation". The default BMR is a 5% change in the group mean relative to the control. Modifications of the response level (i.e., down to 1%; up to $\geq 10\%$) can be made based on the biological significance of the endpoint and other toxicological considerations. The default should not be used when the BMR is specified in other HAS guidelines for a particular endpoint (e.g. cholinesterase inhibition). For comparing the relative toxicity based on the BMD of a same endpoint, the preference is to use ED instead of LED. For comparing the relative sensitivity of multiple endpoints, a "comparable metric" for all endpoints may be desirable, and the BMR of 0.61σ can be used. Modifications of the

¹ USEPA Hill model requires a minimum of 4 data points (3 dose groups plus the control)

² "test 1" of the USEPA BMD software

BMR can be made on the multiplier (e.g., 0.61), based on statistical considerations and the adversity of the endpoint (see: Section IV, Step 4 and Step 6). In the form of the currently available models, BMRs defined in terms of σ should only be applied to normally distributed data³, and their use other than for comparison purposes is not presently recommended.

IV. DISCUSSIONS

This section provides more extensive discussion on the guidance presented in the previous section.

Step 1: Endpoint Selection

In the NOEL-LOEL approach, the critical endpoint selected for a study is the biologically or toxicologically significant effect with the lowest NOEL. The effects with the lowest LED can similarly be selected when using the BMD approach. There is no guarantee, however, that the effect with the lowest NOEL or LOEL also has the lowest ED or LED because these parameters are inherently dependent on the shape of the dose-response curve. Therefore, several endpoints may need to be modeled to determine which has the lowest LED (may or may not be at the same BMR - see Step 6 discussion) and is therefore the critical endpoint. Indiscriminately modeling all endpoints to determine the lowest LED could be burdensome and unnecessary (USEPA, 1995). A more focused approach to ensure capturing the lowest LED is to model endpoints having LOELs up to approximately 5-fold the lowest LOEL. For example, effects having LOELs up to 50 mg/kg/day could be modeled if 10 mg/kg/day was the lowest dose with an established effect for the study. Endpoints without biological or toxicological significance or showing no dose-response relationship can be excluded (USEPA, 2000). For identifying the lowest LOEL or LED from all pertinent inhalation studies, the exposure concentration-duration (e.g., ppm, hours/day) could be expressed in the dose term (e.g., mg/kg/day) to account for the duration variable and species-specific breathing rate. The BMD approach can also be used in the inter-study or inter-species comparison for a specific endpoint.

Unlike the BMD for quantal data, comparisons across continuous endpoints should take into account the different methods for defining BMR (see: Step 5). For example, BMR can be defined in terms of absolute change (e.g., mean RBC counts in 10^6 cells/ μ L) or dimensionless relative change (e.g., mean corpuscular volume, or MCV). Moreover, a given level of BMR (e.g., 5% change) for one endpoint may have a different biological meaning than for other endpoints. Hence, choosing the lowest LED as the critical threshold for risk assessment based on the same level of response across all endpoints may not be valid. The modeler needs to exercise his/her own judgement to determine the BMR level for each endpoint, and take into account the continuum of adverse effects with increasing dose, especially when a sensitive endpoint or biological biomarker is modeled.

Alternatively, the BMD approach can be applied only on a per need basis. For example, when no NOEL can be established in a study (i.e., toxicological effects are observed at the lowest tested dose). In this case, without a BMD approach, the current default for estimating a NOEL

³ The normality of data distribution should be confirmed by a statistical test (e.g. Shapiro-Wilk test for sample size <50) [available from SPSS].

would be to scale down from the LOEL using a somewhat arbitrary uncertainty factor of up to 10. The BMD approach can also be used in the inter-study or inter-species comparisons for specific endpoints.

Step 2: Data Criteria

Both individual data and data expressed in a summarized format (i.e., mean and σ) can be used. Theoretically, individual data allow the model to fully utilize the variable distribution for defining the BMD. The minimum dataset should consist of the dose level, sample size, mean, and a measure of variability (e.g., σ) for each dose group. If data for σ are not available for the exposed groups, the same variance between the control and exposed groups can be assumed, i.e., use the control σ for the exposed groups (USEPA, 2000). This assumption should be clearly stated as an area of uncertainty. However, without σ , a dataset cannot be modeled.

Theoretically, the minimum number of data points for modeling is three. However, more complex models may require more data points. In general, the number of data points cannot fall below the number of estimated parameters and the p value for model fit cannot be calculated when the degrees of freedom is <1 . For example, for the USEPA Hill model, the minimum number of dose groups is 4 for modeling and 5 for assessing the model fit (i.e., calculating the p value). The minimum number of dose groups for a 2-degree polynomial model is 3 for modeling and 4 for a valid model fit assessment (see: Section V for model fit criteria).

In general, datasets with near maximum response at the lowest tested dose would not be a good candidate for modeling because it requires extensive extrapolation to the BMD (see below).

Step 3: Uncertainties in Extrapolation

In general, the closer the response levels of the treatment groups are to the BMR, the less effect the choice of model will have on the estimated ED and LED (USEPA, 1995). Extrapolation beyond the experimentally observable or measurable range is not recommended because different models can yield widely different LED (i.e., greater than a factor of 3) (Crump, 1984; USEPA, 2000). Thus, datasets with near maximum response at the lowest tested dose is generally not a good candidate because it requires extensive extrapolation.

When some degree of extrapolation is necessary (e.g. no NOEL can be determined from a study), consideration is given to the biological plausibility of the shape of the dose-response, especially in the low dose region. For example, in a retrospective epidemiological neurotoxicity study of methylmercury (NAS, 2000), no unexposed individuals were identified, but a threshold was expected to exist in the dose-response metric. Hence, a BMD model (e.g. power model) that allows the dose-response curve to take on a sublinear form for this dataset is considered biologically plausible and appropriate.

Step 4: Data Conversion

Dichotomization may be desirable for comparing the sensitivity (e.g., based on the LED) across multiple endpoints based on the same biological significance level (e.g., 5% probability of

occurrence for the "affected") (USEPA, 2000). The dichotomization method is described here while other treatments of data for defining the BMR are detailed in Step 6.

A reference cutoff is needed to convert (dichotomize) continuous data into a quantal format. The following are two methods to define the cutoff:

A. Explicit Approach

In this approach, continuous data are converted to quantal data with respect to a cutoff. For example, by setting a 300 g cutoff for the body weight reduction, any test subjects having the body weight at or above the cutoff are categorized as "not affected", while those with weights below the cutoff are "affected". A detailed description on the explicit dichotomization procedure was given by Gaylor (1996). One obvious disadvantage of this procedure is the loss of information about the graded magnitude of response.

B. Implicit Approach

In the implicit procedure, the BMR is a specific increase in the proportion of the population falling outside of the probability cutoff. For example, it can be expressed in the form of additional risk:

$$P_{(Y < C | dose_{BMD})} - P_0 = BMR$$

where P_0 : probability an unexposed (control) subject is below a specified response value;

C : background that defines an adverse effect;

$P_{(Y < C | dose)}$: probability an exposed (treated) subject falls below C at the BMD;

Y : any value in the range of response from the test subjects.

The computational procedure for using implicit dichotomization is known as the "hybrid method". Dichotomization is conceptually applied via the risk estimation procedure based on the increased proportion (i.e., probability) of affected individuals above a pre-defined background. In contrast to the loss of information in the explicit method, the implicit modeling approach retains the use of distribution information contained in the continuous data (Crump, 1995). In fact, model simulation studies conducted by Gaylor (1996), West and Kodell (1999), and Crump (2002) indicated that the BMD calculated using the implicit approach has a substantially tighter bound (as indicated by the smaller confidence interval, or, smaller LED/ED ratio) than from the explicit procedure. Hence, when dichotomization is desirable, the implicit method is the method of choice (Gaylor, 1996).

The mathematical procedure for deriving a BMD from continuous data by the implicit method is complex, and requires a computation program. Unfortunately, no software in the public domain is yet available (USEPA, 2000). However, for normally distributed continuous variables, the hybrid modeling approach can be applied indirectly by defining the BMR as a change of the mean response at a specified multiplier of the σ (Crump, 1995). The multiplier is determined based on both the pre-defined background probability (P_0) of "abnormal" (or "affected") individuals in the control population and the excess BMR risk (π) in the exposed population that

is above P_0 . Figure 1 shows the relationship between P_0 , π , and σ . Four sets of the most common combinations of these parameters are given in Table 1 (Crump, 1995). For example, using the USEPA benchmark dose software, the $LED_{0.61\sigma}$ approximates the LED at $\pi = 0.1$, $P_0 = 0.05$. Further details on the criteria of choosing P_0 and π are presented in “Step 6: Define BMR.”

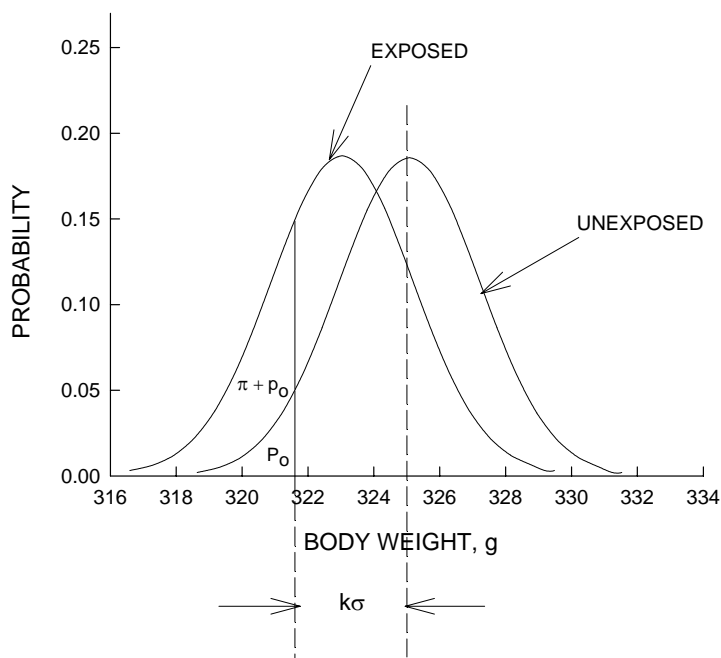


Figure 1. Hypothetical body weight (normal) distribution in an exposed and unexposed experimental animal population (k = a numerical value)

Table 1. Relationship Between P_0 , π , and σ for the hybrid approach (Crump, 1995)

Parameter of Implicit (Hybrid) Method		Multiplier of Standard Deviation (σ)
P_0	π (“BMR”)	k
0.01	0.05	0.77
0.01	0.1	1.1
0.05	0.05	0.36
0.05	0.1	0.61

P_0 = background risk; π = risk above the background; σ : standard deviation; k : multiplier of the standard deviation; to be specified by the user of the USEPA BMD software.

Step 5: Choice of Model and Options

Mathematical Models: The mathematical models available from the USEPA are listed in Table 2.

Model Run Options: The following options are available for the USEPA BMD Models:

- ♦ BMR Type: "relative deviation" from the control (HAS default) , "absolute deviation", $k\sigma$, specified (point) change, and "extra risk" (see Table 3 and the discussion on "Step 6: Define BMR.").
- ♦ Adverse direction: automatic (default), up (increase with dose), or down (decrease with dose).
- ♦ Options for Restriction (coefficient of dose): none (default), non-negative, non-positive. In linear and polynomial models, these selections will determine the restriction on the model coefficient(s), ensuring a strictly decreasing (non-positive) or increasing (non-negative) dose-response curve.
- ♦ Degree of polynomial (n): as the n value increases (i.e., approaching k dose group minus one), the fitted curve may become quite "wavy." (see Section IV)
- ♦ Homogeneous variance ($Rho=0$) or non-homogenous variance ($Rho\neq 0$). This function is to accommodate the dose-dependent change in response variance (as indicated by the failure of the test for homogenous variance). If the $Rho=0$ option is unchecked (i.e., $Rho\neq 0$), dose-dependent variance will be applied by the model for curve fitting (see Appendix A)
- ♦ Restrict $n>1$ (default); n is the shape parameter for characterizing the curvature of the Hill model.
- ♦ Restrict $p>1$ (default); p is the kinetic order of the power model.

Model Fit Criteria: The following criteria should be met for considering a model as adequate to describe the data. Details on the various statistical tests are given in Section V.

- ♦ Model the non-homogeneous variance - allowing the modeling of non-homogenous⁴ variance as a function of dose⁵ (i.e., re-run the model with the option "Rho=0" unchecked in the USEPA BMD software) could improve the BMD model fit⁶.
- ♦ Model goodness-of-fit - χ^2 p value >0.05 (See: model output in Appendix A).
- ♦ Visual examination - inspect the graphical display for the model fit, especially when the goodness of fit p value is not available (Note: p value cannot be calculated when the degrees of freedom is <1).
- ♦ The χ^2 residual values - should not exceed $|2|$ (i.e., absolute value of 2) for each dose group, especially near the BMR.

⁴ "test 2" of the USEPA BMD software

⁵ "test 3" or the 2nd from the last test of the USEPA BMD software

⁶ "test 4" or the last test of the USEPA BMD software

Table 2. Available USEPA BMD Models for Continuous Data

Model	Formula	Application Example
Linear ^a	$\mu(dose) = b + c_1 d$	Body Weight Gain (West & Kodell, 1999)
Polynomial ^a (n=1...k-1 dose groups)	$\mu(dose) = b + c_1 d + c_2 d^2 + \dots + c_n d^n$	RBC Counts (USEPA, 1999)
Power ^a (p=1...k-1 dose groups)	$\mu(dose) = b + s d^p$	Developmental Effects (Allen <i>et al.</i> , 1994a)
Hill ^a (Extension of Power Model)	$\mu(dose) = b + \frac{v d^p}{(d^p + s^p)}$	Ethoxyresorufin-o-deethylase (EROD) response to dioxin Van Birgelen <i>et al.</i> (1995)
Exponential ^b	$\mu(dose) = b [P_b + (1 - P_b) e^{-md}]$	Cholinesterase Inhibition (USEPA, 2002)

Symbols:

$\mu(dose)$: Expected value of the response;

v: Maximum response above background

c_1, c_2, \dots, c_n : Polynomial coefficients;

P_b : Fraction of background activity.

b: Background;

n: Degree of polynomial;

m: Slope-scale factor;

s: Slope;

Power: $p \geq 1$;

d: Dose;

^a/ USEPA (2003) – BMD software version 1.3.2

^b/ USEPA (2002)⁷ – OPcumRisk software

⁷ Used for the Preliminary cumulative risk assessment of the organophosphorus pesticide. The program is set to provide ED and LED at 10% response. Available at <http://www.epa.gov/scipoly/sap/2001/index.htm>

Selection of Model: The best model is selected based on its accuracy in describing the data. Hence, it is recommended that all available models should be run and the model with the best fit would be used. An obvious consideration with the polynomial model is that, while the model fit may appear improved with increasing degree of the polynomial, the dose-response curve becomes “wavy” and lacks scientific support. The model fit for waviness (up and down along the dose axis) usually cannot be prevented by restricting the model parameters to be directional (i.e., non-positive or non-negative, depending on the endpoint parameter).

Considerations for selecting the final model for the BMD analysis are listed below.

- ♦ Consider the biological plausibility of the shape of the dose-response curve, e.g., the methylmercury example in Step 3 above. Avoid over-parameterizing the model (see "Model Run Options" above).
- ♦ Consistency between the model estimated and the observed variables (See: model output in Appendix A), especially at the BMR and BMD region.
- ♦ Use the model estimates with care when the difference between ED and LED is great (i.e., the ED/LED ratio is large, e.g., >5).
- ♦ In general, the model that has the lowest AIC (Akaike's Information Criterion) can be the model of choice (USEPA, 2003; Akaike, 1973; Stone, 1998). However, the selection of a final BMD model among multiple models with adequate fit should not rely solely on the AIC in its current form (Sand *et al.*, 2002). Instead, all of the above considerations should be taken in determining the final model for the BMD analysis.
- ♦ In addition to the above criteria, considerations may be given for using a same model for similar datasets, e.g., data from males and females, or data for a given endpoint from a group of chemicals with the same mode of action.

Improve Model Fit: When the dose-response relationship shows stages of changing slope (e.g., plateau at the high dose range), a better fit at the range near BMR may be achieved by using a more flexible model such as the multistage model with increasing mathematical complexity (i.e., degree of polynomial). However, a simpler model is generally preferred over a more complex one with comparable fit, as model simplicity is included in the AIC calculation (USEPA, 2000). When a full dataset cannot be adequately described with a model (especially if there is a plateau or decreased response with increasing dose), it may be reasonable to focus on the lower dose region, when it is more relevant to the BMR. This can be done by excluding the highest dose data point(s) far above the BMR, while still maintain at least three data points (i.e., controls plus two treatment groups) with a positive trend (i.e., increase in response).

Physiologically based pharmacokinetic (PB/PK) model can also be used to refine the dose estimation for a better model fit. Estimating the dose or concentrations of the parent chemical or its active metabolites at the target site(s) through PB/PK model is especially useful when the plateau of dose-response relationship at the high dose range is due to saturation of metabolic processes or transport systems. When no model can adequately describe the dose response relationship, it may be necessary to revert to the NOEL-LOEL approach.

Step 6: Define BMR

The final step is defining the BMR for which the BMD will be established. Based on the purpose for the BMD modeling, this step defines the metric for measuring the threshold BMR response as well as the quantitative level of the BMR.

Table 3 summarizes the five metrics that are available in the current USEPA BMD software for defining the BMR. They are presented in the order of choice as appeared in the software. The options #1 ("relative deviation" from the control), #2 ("absolute deviation" from the control), and #4 (specified "point" value of response) are conceptually easy to comprehend. The most common form for toxicological response for risk assessment is option #1, expressing the effects as the group mean response relative to the control mean. Options #3 and #5 are discussed under Section C, Endpoint Sensitivity Comparison.

The choice of different metrics, the level of BMR, and the use of ED versus LED may vary depending on the purpose for the BMD analysis. HAS defaults and their rationales are presented for the following three common applications of the BMD approach.

Table 3. Options for Defining the Benchmark Response (BMR) in the USEPA (2003) BMD software.

Definition of BMR	Formula	Examples
1. Relative Change in Mean	$[\mu(d) - \mu(0)] = \delta \cdot \mu(0)$	10% decrease in ChE activities
2. Absolute Change in Mean	$\mu(d) - \mu(0) = \delta$	Decrease in cholinesterase (ChE) activities
3. Change in Mean Relative to Standard Deviation of Control	$[\mu(d) - \mu(0)] = \delta \sigma$	0.61 (i.e., 0.61σ from the control ChE activities)
4. Specified Value (i.e., point)	$\mu(d) = \delta$	Specified ChE activity
5. Change beyond Background Standardized by Total Range of Response (i.e., "Extra" Response).	$[\mu(d) - \mu(0)] / [\mu(\infty) - \mu(0)] = \delta$	Decrease in ChE activities versus net total ChE activities

Key:

$\mu(d)$ and $\mu(0)$ are the mean response at the LED and $d = 0$, respectively.

$\mu(\infty)$ is the limiting (i.e., maximum) mean response as d becomes large; δ is a prescribed value.

A. Threshold for Risk Assessment

For the purpose of characterizing the risk, the BMD is intended to be used as equivalence to the NOEL for deriving the reference dose (RfD) or reference concentration (RfC) (USEPA, 2000).

Background

Using the data from developmental toxicity endpoints (e.g. fetal malformation and body weight change), Kavlock *et al.* (1995) and Allen *et al.* (1994a, b) showed that the LEDs based on a BMR of 5% "relative deviation" from the control were generally similar to the NOELs determined from the same datasets (Table 4). Although the comparison between the BMD and the NOEL has many inherent uncertainties, it presents a reference point for considering a default BMR. A 5% BMR also appears reasonable for endpoints such as body weight changes, since a 10% reduction is considered a marker of toxicity, an indication that the maximum tolerated dose (MTD) has been reached.

Table 4. Comparison of the NOEL and LED derived from the BMD Approach¹.

Endpoints	NOEL/LED ₀₁	NOEL/LED ₀₅	NOEL/LED ₁₀	References
Developmental (% affected fetus/litter – continuous data)	4.3±4.5 (median: 2.5)	1.20±0.88 (median: 0.96)	0.72±0.44 (median: 0.62)	Allen <i>et al.</i> , 1994a ²
Developmental (% affected fetus/litter – continuous data)	-	0.78 (range: 0.42-1.67)	0.4 (range: 0.21-0.83)	Kavlock <i>et al.</i> , 1995 ³

^{1/} Data were expressed as the ratio between the NOEL and the LED.

^{2/} The NOELs were based either on expert judgment or iterative trend test (removing the highest data point until no significant trend was present). Endpoints included fetal death and gross, visceral, and skeletal malformations in mice, rabbits, rats, or hamsters. The continuous power model was used for the continuous data.

^{3/} The NOELs were determined as above. Fetal weight change was the endpoint. The continuous power model was used for the continuous data.

HAS guidance

The HAS default BMD is the LED₀₅, the LED at 5% BMR based on the "relative deviation" (i.e., option #1 of USEPA BMD software BMR matrices). For example, the BMD for the body weight endpoint would be the LED for 5% mean body weight reduction over the mean control weight. This default is consistent with the aforementioned studies by Kavlock *et al.* (1995) and Allen *et*

al. (1994a, b), and supported by the definitions of “adversity” at a higher percentage of response, e.g. MTD.

Flexibility is given for scaling upward ($\geq 10\%$) and downward (1%) from the default 5% BMR based on the following considerations. The guidance given for quantal data in DPR MT-1 (2004) should be consulted for a more detailed discussion on this subject.

- **Biological significance of the endpoint.** The general guide for endpoint severity given for quantal data (Table 3 in DPR MT-1, 2004) can also be used for continuous data. A lower BMR (e.g., 1%) can be justified for more detrimental effects and a higher BMR (e.g., above 10%) can be justified for milder effects or sensitive biomarkers.
- **Additional auxiliary toxicity data.** For example, Woutersen *et al.* (2001) suggested that the BMR could be 30% for elevated AST activity because of its wide variability and limited implication as a biomarker for hepatocellular damage. However, a lower BMR (e.g., 10%) for AST activity may be justified if liver necrosis is also observed in the same animals.
- **Technical limitation in endpoint detection.** This was the reason for the BMR of 10% decrease in brain cholinesterase (ChE) activity used in assessing the cumulative risk of organophosphate insecticides (USEPA, 2001).

It should be noted that, the general default BMR presented in this document should not be used when the BMR is specified in other HAS guidelines for a particular endpoint (e.g. cholinesterase inhibition).

B. Relative Toxicity

Background

The ratio of two BMDs at a given BMR of an endpoint is often used as a measure of relative toxicity in expressing the gender and species sensitivity to a chemical. The concept of relative toxicity is also applicable for assessing the risk of exposure to a mixture of chemicals with the same mode of action. For example, in assessing the risk of exposure to multiple organophosphate (OP) pesticides, USEPA estimated the Relative Potency Factors (RPFs) of more than 20 OPs based on the ratio of their BMDs to the BMD of an index OP chemical (in this case, methaminophos) at 10% brain cholinesterase inhibition in female rats (USEPA, 2002). These RPFs were then used to scale and sum the exposure from all OPs for calculating the overall margin of exposure (MOE)⁸.

HAS guidance

When comparing the sensitivity of a specific endpoint among studies, species, or chemicals at a specified BMR, it is desirable to base the comparison on the ED rather than the LED. The use of the ED avoids the uncertainty associated with the model-dependent tendency in the LED

⁸ Margin of Exposure is the ratio of the toxicity threshold (e.g., NOEL, BMD) to the exposure.

estimation, especially when more than one model is used to estimate the BMDs for the entire database.

C. Endpoint Sensitivity Comparison

In risk assessment, the most sensitive endpoint is often defined as the endpoint that has the lowest threshold. In the NOEL-LOEL approach, this is usually the endpoint with the lowest NOEL. For the BMD approach, this is the endpoint with the lowest LED at a pre-determined BMR. However, unlike for the quantal data, the different options for the BMR metrics present a greater complexity for the continuous data. As the quantal response represents a discrete measure of outcome (i.e., binary), its BMR can be expressed in terms of probability of response, e.g., "extra risk" $([P(d)-P(0)]/[1-P(0)])$ or "additional risk: $(P(d)-P(0))$ " (see DPR MT-1, 2004; the Guidance for Quantal Data). Thus, by defining the BMR on the same probability term, the BMD may be used as an indicator of sensitivity across all quantal endpoints (e.g., lower BMD signifies greater sensitivity). On the contrary, continuous data can assume any value in a range of measures (e.g. body weight = 100 to 400g) and can also be expressed in terms of their changes (e.g. body weight gain). Thus, a 5% BMR for the continuous endpoint does not have the same meaning for all continuous endpoints. Nor does it carry the same meaning to quantal endpoints at the same numeric value of BMR.

Background

Presented below are the various proposals for an objective "comparable metric" among all endpoints, and the limitations in their current forms:

- ◆ Use explicit dichotomization of data as described in Step 4. The loss of useful data through this process tends to "overestimate" the lower bound of the BMR, i.e., yields overly "conservative" (low) LED (Gaylor, 1996, Kodell and West, 1999, Crump, 2002).
- ◆ Standardize the change beyond the background by the total range of response (Option 5 in Table 3). The form of $[\mu(d)-\mu(0)]/[\mu(\infty)-\mu(0)]$ (BMR Option #5 of the USEPA BMD software) appears to be functionally similar to that of the extra risk for the quantal data (i.e., $[P(d)-P(0)]/[1-P(0)]$) in that, both definitions involve a normalization of the "response". Based on their similarity, Murrell *et al.* (1998) suggested that $[\mu(d)-\mu(0)]/[\mu(\infty)-\mu(0)]$ (where $\mu(\infty)$ is the limiting ("maximum") response) could serve as a common basis for comparing BMDs derived from different continuous endpoints. Unfortunately, this metric has a rather limited application (Crump, 2002). For example, when little or no information is available in the data for estimating $\mu(\infty)$, the denominator $[\mu(\infty)-\mu(0)]$ becomes either undefined or highly uncertain. Accordingly, the application of the method to characterize a toxicological endpoint would also become impossible or erroneous.
- ◆ Define the change relative to the standard deviation of the response in the unexposed (i.e., control) subjects. For continuous data with a constant variance ($\text{Rho}=0^9$), this is

⁹ The meaning of the hybrid method output is uncertain when the variance from all test groups are not homogeneous.

equivalent to the hybrid method described in Step 4 (Gaylor and Slikker, 1990; Kodell and West, 1993; West and Kodell, 1993; Crump, 1995).

For the hybrid method, any multiplier of σ can theoretically be used to generate the BMDs for comparing the sensitivity of endpoints, as long as the same multiplier is used for all endpoints. However, there may be some advantage for a default multiplier that considers the conventional and practical use of P_0 and π . In the study by Kavlock *et al.* (1995), the LED for a 0.5σ BMR for fetal body weight was numerically similar to the developmental NOEL. Using the hybrid method, Crump (1995) demonstrated that defining the BMR of $P_0 = 0.05$ and $\pi = 0.1$ (Table 1) is equivalent to 0.61σ for normally distributed continuous variables. The P_0 value is consistent with the definition of abnormal range in clinical data (i.e., $P_0 = 0.05$ means beyond the 95th percentile of a normal distribution) (Crump *et al.*, 1995). The π value is consistent with the quantal BMR at which Allen *et al.* (1994) reported that the QLED₁₀ (the LED₁₀ for quantal data) was approximately the same as the NOEL for developmental toxicity endpoints. Based on the results of Kavlock *et al.* (1995) and Crump *et al.* (1995), Haber *et al.* (1998) recommended that, in the absence of specific information for defining the BMR, the combination of $P_0 = 0.05$ and $\pi = 0.1$ can be used as a default. In a similar proposal, 1σ was recommended by the USEPA. As shown in Table 1, this is equivalent to the BMR at $P_0 = 0.01$ and $\pi = 0.1$ of the hybrid method. Although the choice of lower P_0 (i.e., $P_0 = 0.01$, instead of 0.05 as proposed by Haber *et al.*, 1998) is consistent with the notion that a more extreme (i.e., lower P_0) background probability means greater certainty of adverse effects, it also means that a higher dose (i.e., less conservative) is needed to cause the same percent increase in risk (Figure 1).

As discussed above, hybrid method allows the sensitivity of continuous endpoints to be compared on an equal statistical basis. However, an extreme caution needs to be exercised when comparing BMDs derived from the hybrid method to the conventionally determined NOELs and LOELs. In the hybrid method, the BMD is estimated based on the change in the probability of an individual adverse response. For example, $\pi = 10\%$ for a “lower” body weight means that 10% of the individuals in a population have a significant body weight reduction beyond the cutoff P_0 value (e.g., 95th percentile; $P_0 = 0.05$). This definition of BMR is different from both the conventional criteria for defining the NOEL-LOEL of continuous data, and the common BMR definitions (i.e., Options #1, 2, and 4 in Table 3) based on the changes in population mean (i.e., P_0 at 0.5). Thus, the BMD defined by the hybrid method (e.g. 0.61σ) does not have the same quantitative meaning as those established from either the continuous NOEL-LOEL approach or the “non-hybrid” BMD modeling.

Although hybrid models appear to be favored by some investigators (Harber *et al.*, 1998; West and Kodell, 1999; Budtz-Jørgensen *et al.*, 2000; Crump *et al.*, 2000; NAS, 2000), the computational procedure is very complex and a computer program needed for this calculation is currently not available (USEPA, 2000). Apart from the aforementioned disparity to the conventional BMR expression, there are additional limitations with the hybrid method proposed by Crump (1995). In the form that is used in the currently available models, the method is limited to normally distributed data with homogeneous variances. Depending on the data characteristics, defining the BMR in terms of σ could also result in a higher LED with a greater σ , having the appearance of “rewarding” a bad study. Thus, other than for comparison purposes,

a “stand alone” application of the hybrid method is not recommended for establishing the BMD for risk assessment.

HAS guidance

When an “equivalent metric” is desirable, a BMR of 0.61σ can be used. The multiplier of 0.61 can be modified by statistical and biological considerations (i.e., a different combination of P_0 and π values as in Table 1). It should be emphasized that, in the form used in currently available models, this approach should only be used for distribution data with homogeneous variances. Statistical tests (e.g., Shapiro-Wilk test) for normality should be performed to confirm the assumption of normal distribution. A “stand alone” use (i.e., other than for comparison purpose) of 0.61σ for defining the BMR is not recommended especially when a “larger” standard deviation is associated with a poorer study quality.

V. MODEL OUTPUT

A sample text and graphic output for a 10% "relative deviation" by the Hill model is given in Appendix A. It provides information on the parameter estimates, the statistical tests, and the “BMD” (e.g., ED_{10}) and “BMDL” (e.g., LED_{10}). Explanations of the corresponding statistical tests are presented below. The Help manual (USEPA, 2003) should be consulted for further information.

χ^2 Residual: The χ^2 residual values for each dose group given the "Table of Data and Estimated Values of Interest" should not exceed $|2|$ (i.e., absolute value of 2), especially for the low dose groups.

Test 1: Difference Among Dose Groups: Determines whether there is a significant ($p < 0.05$) difference among the responses and/or variances for the dataset. The dataset must pass this test to be modeled.

Test 2: Homogeneity of Variance: Determines whether the variances (squares of the standard deviations) are significantly ($p > 0.05$) different among the dose groups. If they are judged as "non-homogeneous" (test 2, Appendix A), the model is re-run with the “constant variance; $Rho=0$ ” unchecked in the run screen. If the variance is homogeneous as indicated by test 2, the next test (test 3) is the final test (for the BMD model fit).

Test 3: Modeling of Variance: Determines whether the non-homogeneous variances (as indicated in test 2) can be modeled. Unchecking the option " $Rho=0$ " allows the modeling of the dose-variance relationship. The test for appropriateness of the variance modeling is given in test 3. Since the power function is the only model in the current BMD software, it might not be applicable for all datasets. If the power function is not appropriate (i.e., failing test 3), the BMD analysis can be re-modeled with the assumption of homogeneous variance (re-run the model with " $Rho=0$ " checked).

Test 4: Goodness of Fit: Determines whether the particular selected model adequately describes the data (χ^2 goodness of fit). This test is model specific and needs to be examined for every model and every new set of model parameters. Models that do not

pass the goodness of fit test ($p < 0.05$) should not be used for estimating the BMD. If the goodness-of-fit test cannot be calculated due to zero degrees of freedom, visual inspection and χ^2 residuals can be used to determine whether the model provides a good fit.

VI. REFERENCES

- Akaike, H. 1973. "Information theory and an extension of the maximum likelihood principle." In: Proceedings of the Second International Symposium on Information Theory, B. N. Petrov and F. Csaki, Eds. Akademiai Kiado, Budapest, pp. 267-281.
- Allen, B., C. Kimmel, R. Kavlock, and E. Faustman. 1994a. Dose-response assessment for development toxicity II. Comparison of generic benchmark dose estimates with NOAELs. *Fund. Appl. Toxicol.* 23:487-495.
- Allen, B. C., R. J. Kavlock, C. A. Kimmel, and E. M. Faustman, E.M. 1994b. Dose-response assessment for development toxicity III. Statistical models. *Fund. Appl. Toxicol.* 23:496-509.
- Budtz-Jorgensen, E., P. Grandjean, N. Keiding, R. F. White, and P. Weihe. 2000. Benchmark dose of methylmercury-associated neurobehavioral deficits. *Toxicology Lett.* 112-113:193-199.
- Crump, K. S. 1984. A new method for determining allowable daily intakes. *Fund. Appl. Toxicol.* 4:854-871.
- Crump, K. S. 1995. Calculation of benchmark doses from continuous data. *Risk Anal.* 15:79-89.
- Crump, K. S. 2002. Critical issues in benchmark calculations from continuous data. *Critical Rev. Toxicol.* 32:133-153.
- Crump, K. S., J. Viren, A. Silver, J. Clewell, J. Gearhart, and A. Shipp. 1995. Reanalysis of dose-response data from the Iraqi methylmercury poisoning episode. *Risk Anal.* 15:523-532.
- Crump, K. S., C. V. Landingham, C. Shamlaye, C. Cox, P. W. Davidson, G. J. Myers, and T. W. Clarkson. 2000. Benchmark concentrations for methylmercury obtained from the Seychelles child developmental study. *Environ. Health Persp.* 108:257-263.
- DPR MT-1, 2004. Guidance for Benchmark Dose (BMD) Approach - Quantal data. Medical Toxicology Branch, Department of Pesticide Regulation, California Environmental Protection Agency, Sacramento, CA.
- Gaylor, D. W. 1996. Quantalization of continuous data for benchmark dose estimation. *Regul. Toxicol. Pharmacol.* 24:246-250.
- Gaylor, D. W. and W. Slikker. 1990. Risk assessment for neurotoxic effects. *NeuroToxicology* 11:211-218.
- Gaylor D, L. Ryan, D. Krewski, and Y. Zhu. 1998. Procedures for calculating benchmark doses for health risk assessment. *Regul. Toxicol. Pharmacol.* 28:150-164.

- Haber, L.T., B. C. Allen, and C. A. Kimmel. 1998. Non-cancer risk assessment for nickel compounds: Issues associated with dose-response modeling of inhalation and oral exposures. *Toxicol. Sci.* 43:213-229.
- Kavlock, R. J., B. C. Allen, E. M. Faustman, and C. A. Kimmel. 1995. Dose-response assessment for development toxicity IV. Benchmark dose for fetal weight changes. *Fund. Appl. Toxicol.* 26:211-222.
- Kodell, R. L. and R. W. West. 1993. Upper confidence limits on excess risk for quantitative response. *Risk Anal.* 13:177-182.
- Murrell, J. A., C. J. Portier, and R. W. Morris. 1998. Characterizing dose-response I: Critical assessment for the benchmark dose concept. *Risk Anal.* 18:1326.
- NAS (National Academy of Sciences). 2000. Toxicological Effects of Methylmercury. National Academic Press. Washington, D.C.
- Sand, S., A. F. Filiposson, and K. Victorin. Evaluation of the benchmark dose method for dichotomous data: Model dependence and model selection. *Regul. Toxicol. Pharmacol.* 36(2):184-197, 2002
- Stone, M. 1998. Akaike's Criteria. In: Encyclopedia of Biostatistics. Volume 1, A-COX. Armitage, P. and T. Colton, eds. John Wiley and Sons, N.Y.
- USEPA. 1999. Integrated Risk Information Service (IRIS). Integrated Risk Information Service (IRIS) U.S. Environmental Protection Agency. Cincinnati, Ohio.
- USEPA. 2000. Benchmark Dose Technical Guidance Document. EPA/630/R-00/001 (draft)
- USEPA. 2002. Preliminary cumulative risk assessment of the organophosphorous pesticide. U.S. Environmental Protection Agency.
- USEPA. 2003. Help Manual for Benchmark Dose Software Version 1.3.2. Office of Research and Development, Washington D.C.
- Van Birgelen, A. P. J. M., J. Van Der Kolk, K. M. Fase, I. Bol, H. Poiger, A. Brouwer, and M. Van Den Berg. 1995. Subchronic dose-response study of 2,3,7,8-tetrachlorodibenzo-p-dioxin in female Sprague-Dawley rats. *Toxicol. Appl. Pharmacol.* 132:1-13.
- West, R. W. and R. L. Kodell. 1993. Statistical methods of risk assessment for continuous variables. *Commun. Statist-Theory Meth.* 22:3363-3376.
- West, R.W. and R. L. Kodell. 1999. A comparison of methods of benchmark dose estimation for continuous response data. *Risk Anal.* 19, 453-459.

Woutersen, R. A., D. Jonker, H. Stevenson, J. D. te Biesebeek, and W. Slob. 2001. The benchmark dose approach applied to a 28-day toxicity study with rhodorsil silane in rats: the impact of increasing the number of dose groups. *Fd. Chem. Toxicol.* 39:697-707.

Appendix A

Example Output from USEPA BMD Program

```
=====
Hill Model. $Revision: 2.1 $ $Date: 2000/10/11 21:21:23 $
Input Data File: H:\BMD\DATA\XXXXXXXXXXXX_M_R_12_MO.(d)
Gnuplot Plotting File: H:\BMD\DATA\ XXXXXXXXXXXXXX _M_R_12_MO.plt
Fri Apr 25 10:38:23 2003
=====
```

BMDS MODEL RUN

The form of the response function is:

$$Y[\text{dose}] = \text{intercept} + v \cdot \text{dose}^n / (k^n + \text{dose}^n)$$

Dependent variable = MEAN

Independent variable = Dose

Power parameter restricted to be greater than 1

The variance is to be modeled as $\text{Var}(i) = \alpha * \text{mean}(i) ^ \rho$

Total number of dose groups = 5

Total number of records with missing values = 0

Maximum number of iterations = 250

Relative Function Convergence has been set to: 1e-008

Parameter Convergence has been set to: 1e-008

Default Initial Parameter Values

```
alpha = 0.00548995
rho = 1.20547
intercept = 0.74
v = -0.62
n = 0.978102
k = 0.4
```

Asymptotic Correlation Matrix of Parameter Estimates

(*** The model parameter(s) -n have been estimated at a boundary point,
or have been specified by the user, and do not appear in the
correlation matrix)

	alpha	rho	intercept	v	k
alpha	1	0.84	-0.12	0.04	0.12
rho	0.84	1	-0.12	0.038	0.14
intercept	-0.12	-0.12	1	-0.55	-0.67
v	0.04	0.038	-0.55	1	-0.18
k	0.12	0.14	-0.67	-0.18	1

Parameter Estimates

Variable	Estimate	Std. Err.
alpha	0.00526837	0.00196084
rho	1.21224	0.316443
intercept	0.731768	0.0164364
v	-0.731088	0.0179037
n	1	NA
k	0.532802	0.0600146

NA - Indicates that this parameter has hit a bound implied by some inequality constraint and thus has no standard error.

Table of Data and Estimated Values of Interest

Dose	N	Obs Mean	Obs Std Dev	Est Mean	Est Std Dev	Chi^2 Res.
-----	---	-----	-----	-----	-----	-----
0	10	0.74	0.07	0.732	0.0601	0.137
0.1	10	0.61	0.05	0.616	0.0541	-0.115
0.3	10	0.46	0.04	0.468	0.0458	-0.183
0.9	10	0.28	0.04	0.273	0.033	0.226
2.7	10	0.12	0.02	0.121	0.0202	-0.058

Model Descriptions for likelihoods calculated

Model A1: $Y_{ij} = \mu(i) + e(ij)$
 $\text{Var}\{e(ij)\} = \sigma^2$

Model A2: $Y_{ij} = \mu(i) + e(ij)$
 $\text{Var}\{e(ij)\} = \sigma(i)^2$

Model A3: $Y_{ij} = \mu(i) + e(ij)$
 $\text{Var}\{e(ij)\} = \alpha * (\mu(i))^\rho$

Model R: $Y_i = \mu + e(i)$
 $\text{Var}\{e(i)\} = \sigma^2$

Likelihoods of Interest

Model	Log(likelihood)	DF	AIC
A1	130.616461	6	-249.232922
A2	137.681683	10	-255.363365
A3	136.981376	7	-259.962753
fitted	136.246742	5	-262.493484
R	48.680356	2	-93.360713

Explanation of Tests

Test 1: Does response and/or variances differ among Dose levels? (A2 vs. R)
 Test 2: Are Variances Homogeneous? (A1 vs A2)
 Test 3: Are variances adequately modeled? (A2 vs. A3)
 Test 4: Does the Model for the Mean Fit? (A3 vs. fitted)

Tests of Interest

Test	-2*log(Likelihood Ratio)	Test df	p-value
Test 1	178.003	8	<.0001
Test 2	14.1304	4	0.00689
Test 3	1.40061	3	0.7054
Test 4	1.46927	1	0.2255

The p-value for Test 1 is less than .05. There appears to be a difference between response and/or variances among the dose levels It seems appropriate to model the data

The p-value for Test 2 is less than .05. A non-homogeneous variance model appears to be appropriate

The p-value for Test 3 is greater than .05. The modeled variance appears to be appropriate here

The p-value for Test 4 is greater than .05. The model chosen seems to adequately describe the data

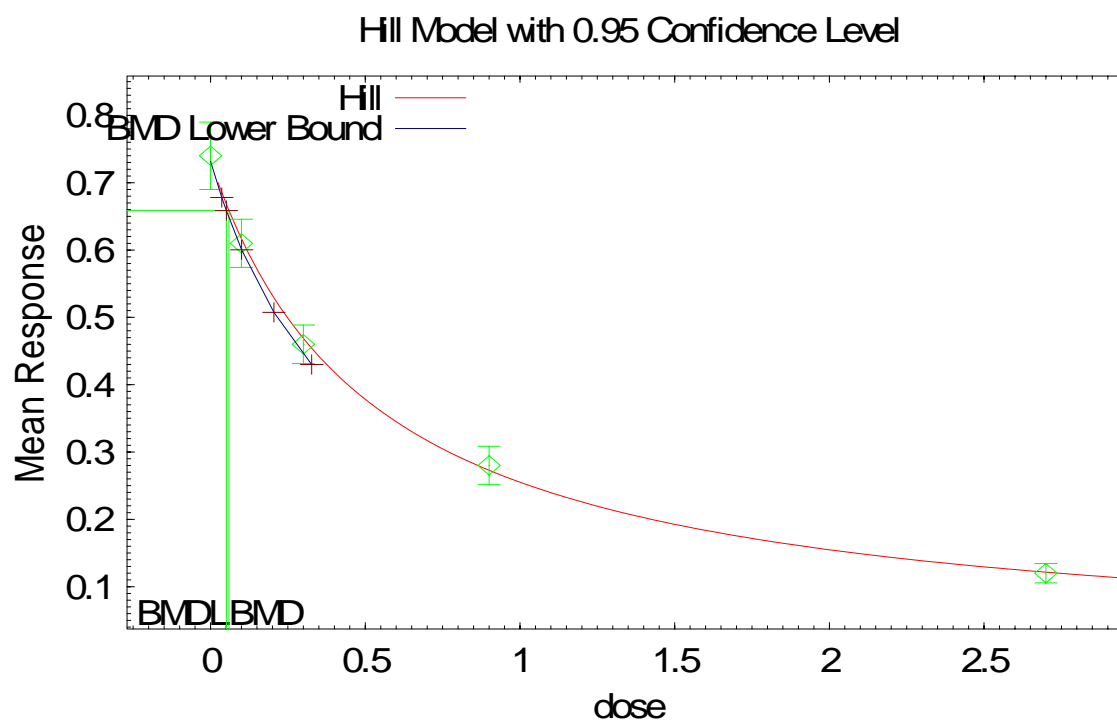
```
Benchmark Dose Computation
Specified effect =          0.1

Risk Type          =      Relative risk

Confidence level =          0.95

      BMD =      0.0592614

      BMDL =      0.050923
```



10:38 04/25 2003